DOCUMENT RESUME

ED 242 713 SP 024 132

AUTHOR Danzig, Arnold B.
TITLE Teacher Testing as a Condition of Certification: Promises, Pitfalls, and Pratfalls.
PUB DATE May 83
NOTE 33p.
PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Standards; *Basic Skills; Beginning Teachers; College Admission; Competency Based Teacher Education; *Court Litigation; Criterion Referenced Tests; *Minimum Competency Testing; Minority Groups; Screening Tests; Standardized Tests; State Standards; *Teacher Certification; *Test Validity
IDENTIFIERS *Arizona

ABSTRACT

This paper examines some impacts of basic skills proficiency testing of teachers in Arizona. A discussion of some legal issues involved in teacher proficiency testing is presented. Specific cases are cited which illustrate the fact that the courts have ruled that testing must show "job-relatedness" and a "rational basis for use." It is pointed out that, in spite of high rates of minority failure, the courts have ruled that basic skills proficiency testing is related to the states' "legitimate employment needs." It is noted that court rulings on test validation have steered states toward criterion-referenced examinations which measure basic skills, professional knowledge, and content area ability. The Arizona Performance-Based Teacher Certification program, which consists of a basic skills subtest, a professional knowledge subtest, and a proposed competency-based two year teacher residency program, is described. The major focus of this description is on the Basic Skills subtests and some impacts this has had on teacher education in Arizona. Results of the Arizona Teacher Proficiency Exam in the basic skills component statewide, and at Northern Arizona University, are examined. Finally, some early testing results are analyzed. This includes a discussion of some test score correlations with other measures of student ability as well as a discussion of implications for minority students. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED242713

TEACHER TESTING AS A CONDITION OF CERTIFICATION: PROMISES, PITFALLS, AND PRATFALLS

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Arnold B. Danzig
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Arnold B. Danzig
Assistant Professor
Educational Administration,
Supervision and Foundations
Northern Arizona University

May, 1983

SP 024 132

TEACHER TESTING AS A CONDITION OF CERTIFICATION: PROMISES, PITFALLS, AND PRATFALLS

Arnold B. Danzig
Northern Arizona University

INTRODUCTION

One might argue that there exists today a crisis of confidence in teacher education. Some have argued that the teacher job market collapse of the 1970s played an important part in shrinking the available talent pool (Weaver, 1979). Fewer job openings and the growth of opportunities and salaries in other fields, particularly for women, have led to a deteriorated applicant pool.

One indication of decline that is often pointed to comes from the test results of education majors on a variety of standardized measuring instruments including the Scholastic Aptitude Test (SAT), Graduate Record Examination (GRE), and the National Teachers Examination (NTE) (see Weaver, 1979; Hathaway, 1980; Ivie, 1982). According to Weaver (1979) education majors score near the bottom on SAT Math and English Tests, surpassing only those in Office-Clerical and Vocational-Technical fields.

Responses to this perceived decline have come from educators, the media, parents, the military, business leaders, voicing their concern over the outcomes of public education. In turn, state legislatures, colleges of education, and departments of education have begun to question and readdress teacher education in the various states across the United States.

The legal responsibility for public education has traditionally resided with state governments. Under most state constitutions, the state legislature is authorized to provide public education in consort with the local district.

As legislatures have become more concerned with the quality of education existing in the schools, they have passed legislation requiring new programs and teacher credentialing standards. Requiring competency testing for teachers in basic skill areas (reading, grammar, math) as well as subject areas has proven to be an expedient way for legislatures to deal with the complex problem of improving public education.

Rubinstein et al. (1982) report that as of January, 1982, sixteen states have newly revised programs for teacher certification in process or under consideration. New programs exist in Florida, Georgia, Louisiana, Mississippi, Virginia, and West Virginia. In 1981-1982 alone, eight other states have implemented major program changes while California, New York, and Texas are currently planning implementation of new programs.

Elements of Teacher Testing Programs

State programs vary in the number and nature of tests used, but most involve some combination of three basic elements: (1) Basic Skills testing; (2) Professional Studies Testing; and (3) Content Area Testing. Five states use or intend to use all three types of examinations: Arkansas, New York, Mississippi, Virginia, and West Virginia. Six others use or will use two of the above procedures: Alabama, Arizona,

Florida, North Carolina, and Tennessee (Rubinstein et al., 1982).

In addition, some programs include a supervised teacher internship program as a condition of certification. Eight states have, or project to have an internship ranging from one to three years in length: Alabama, Arizona, Florida, Georgia, Oklahoma, North Carolina, New York, and South Carolina (Kelley, 1982; Rubinstein et all, 1982).

The National Teachers Examination

The most commonly used instrument employed by states involved with teacher certification testing has been the National Teachers Examination (NTE), used for a variety of purposes in twenty-five states (Ivie, 1982). The NTE is a norm-referenced test prepared by Educational Testing Service of Princeton, New Jersey. It consists of a Common Examination made up of 270 professional and general education questions to which three hours and fifteen minutes are allotted and which costs the examinee a fee of $35.00 (Pearson, 1981). A maximum score of 990 is possible on the Common Exam. An additional two hours, 150 question Subject Area examination is also part of the NTE making for a total possible score of 1,980.

Some states (Louisiana and South Carolina) have recommended cut-off scores on the NTE as a condition of certification. However, critics have argued that teacher competence is not a unitary trait to be measured by a paper and pencil multiple choice test, but a process consisting of a repertoire

of skills (Medley, 1981; Kelley, 1982).

Controversy over use of the NTE has centered on two issues: reliability of the instrument to measure teacher competence and high failure rate of minorities. Ivie (1982) found that 96 percent of students at three black institutions in Georgia had NTE scores below the fiftieth percentile compared to 49 percent of students at eleven predominantly white institutions. Wilson and Stoller (1981) found little criterion-related validity to support the use of the NTE science related tests. Meritt (1980) on the other hand, found NTE Professional education test scores to be positively correlated to grades achieved in certain teacher education courses.

Legal issues in Teacher Testing

The courts have generally ruled that if tests are to be used for screening out applicants, besides adhering to due process law requirements, safeguarding property rights by procedural fairness, and requirements against capricious use, tests must also abide by EEOC guidelines to meet certain standards of validity and specific job-relatedness (Rubinstein et al., 1982).

EEOC guidelines specify methods of validating tests, and stipulate that differential failure rates must have a job relevant basis. Violations of job-relatedness and validation procedures were heard in Watson v. County School Board of Nansemond County, and Albemarle Paper Company v. Moody (Hathaway, 1980; Nassif & Solomon, 1982).

A Virginia court ruled in the case of Watson, that the NTE "as it was used" for purposes of retention of experienced teachers or dismissal of teachers was inappropriate, since neither detailed job analysis descriptions, nor validation studies had been performed. The court in Albemarle, referring to the importance of analyzing the "attributes and particular skills" needed for a job, established clear criteria to determine whether employers' tests were in fact, job related (Rubinstein et al., 1982).

Rationality Relationships

Tests and certification requirements must also reflect a rational basis for use.

In Dent v. West Virginia, the court ruled that certification qualification for doctors, must "bear a direct relationship" and be "appropriate to a particular profession or calling" i.e., show some degree of "rationality" or a show that a rational relationship exists.

In Chance v. Board of Examiners of City of New York, New York licensing requirements, for school supervisory personnel were declared invalid, because they lacked job relevance. A disparate effect had been found on the licensing exam, which lacked content validity, among black and Puerto Rican school principals.

Rational relationship is further illustrated in Armstead v. Starkville Municiple Separate School District, Tyler v. Vickery, and Georgia Association of Educators, Inc. v. New York. In Armstead, use of GRE scores for retention

or hiring of school personnel was deemed "irrational." However, in Tyler, the court ruled that the Georgia Bar Exam was not discriminatory, since it had a rational bearing, accepted validity, and purposeful design. The Georgia Association of Educations struck down the use of NTE scores for granting six-year certification, based on ETS testimony that the test was designed to measure academic knowledge, not competence of teachers, and that arbitrary cut-off scores could not be validated (Hazard et al., 1977).

Discriminatory Intent

The court ruled in Washington v. Davis that a written test for police applicants in Washington, D.C. was in constitutional violation, since it had a disproportionate black failure rate, and was not shown to measure actual job performance. Later, the Supreme Court reversed the decision by claiming it could not find a racial classification for the test, and that the test held a "rational relationship" to job performance without evidence of the intent to discriminate (Hazard et al., 1977).

Business Necessity

The requirement to show "business necessity" relationship is most widely noted in Griggs v. Duke Power Company where an employer was allowed to use employment tests having a disproportionate impact on minority applicants if the employer could show an absence of "discriminatory intent" and show a "manifest relationship" between the knowledge required on the test and the employment.

An Emerging Model of Teacher Certification

Based on concerns and aspirations voiced by students and teachers, college faculty and state departments of education, parents and legislators, teacher unions and the courts, an emerging consensus for teacher education and certification may be at hand. Such a program consists of at least three basic elements: (1) Competency-based teacher education program; (2) job-related written examinations as a condition of certification; and (3) active feedback and evaluation of instruction to include in some cases, an internship.

Competency-based teacher education attempts to define a clear criteria for teaching which can be tested and evaluated at a later date (Piper & Houston, 1980). Elam (1971) defines CTBE as bringing about learner achievement of publicly stated performance goals or competencies and attaining evidence of learner outcomes by assessing performance under specified condition. The learner is guided through measureable outcomes with emphasis placed on professional practice in the classroom.

Teacher testing in basic skills, professional knowledge, and subject expertise as a measure of quality and performance seems to be a favored way of providing for competent teacher candidates. Ninety-five percent of people responding to a recent Gallop Poll believed that teachers should be required to pass an examination in their subject area (Vorwerk & Gorth, 1982). Ornstein (1981) reports that every state will enact some form of teacher accountability over the next decade.

Increased emphasis with on-the-job performance, concern with performance objectives, and dissatisfaction with the norm referenced National Teacher Examination, has led to the development of criterion-reference tests. Emphasis is placed on identification of specific learning objectives, knowledge and skills, determined in advance and measureable. Legal requirements for objective-based and job relevant methods of measurement also seem to be more adequately satisfied by such an approach.

An example of the third aspect of teacher certification -- active feedback on instruction -- can be seen in the following example. Brickell (1980) reports that New Jersey's minimal skills testing program centers on a five person team of professionals who visit schools where students fail to meet testing standards during three consecutive years. Recommendations to improve learning and instruction are made following research-based standards for good instruction provided by the NIE Regional Laboratory, Research for Better Schools. Four state educational improvement centers provide technical assistance for local schools to close the failure gap.

The next section follows the Arizona experience more closely as a number of possible directions and strategies are explored.

Teacher Certification in Arizona

During the past few years, the Arizona Department of Education, responded to legislative mandate by embracing a comprehensive program to systematically evaluate new teachers.

Known as the Arizona Performance-Based Teacher Certification program, it consists of a basic skills subtest, a professional knowledge subtest, and a proposed competency based two year teacher residency program.

The major focus of this analysis is on the Basic Skills subtests and some of the impacts this has had on teacher education in Arizona. However, before addressing this issue in detail, it is worth describing the total program in general.

The Arizona legislature now requires that all new teachers, or teachers new to the State, pass a proficiency examination in reading, grammar, and mathematics as a condition of certification. The responsibility for test selection, administration, and determination of passing scores was given to the Arizona Department of Education.

ATPE: Basic Skills Component

The State moved from legislative directive in April, 1980, to full implementation of the Arizona Teacher Proficiency Exam (ATPE) in June, 1981, through several stages. To meet the legal implementation date of October, 1980, the Common Examination of the National Teachers Exam was used on an emergency basis. A task force set up by the Department of Education then evaluated previously validated test items in reading, grammar, and mathematics from a test bank developed through the Los Angeles County Superintendent's Office and determined the content appropriate for use in the ATPE basic skills component.

The ATPE is comprised of 150 questions, with 50 each

in reading, grammar, and mathematics. An arbitrary passing score of fifty percent correct responses was established during the six month norming period, to be phased upward in two stages, so that by January, 1982, eighty percent correct responses was required for a passing score.

ATPE: Professional Knowledge Component

In December, 1980, the Department of Education received a State directive to develop a professional education subtest to measure knowledge of "essential skills." With the assistance of five task teams, the Department assessed what it considered to be the specific and essential professional skills that should be learned in teacher education programs. The result was a list of 73 necessary skills from six content areas: classroom management, curriculum and instruction, assessment and evaluation, growing and learning theories, educational foundations, and organization and administration. Tests items were designed to be given as the Professional Knowledge subtest of the ATPE. Plans to raise the required correct response level to eighty percent have for the moment been postponed, and at the present time, fifty percent of the questions on this subtest must be answered correctly to achieve a passing score.

Arizona Teacher Residency Program (ATRP)

The third and final phase of the Arizona Performance-Based Teacher Certification Program authorized by state law, is a proposed two-year teacher residency program presently in its pilot testing phase.

ATRP participants will have two years to demonstrate minimal levels of acceptable classroom performance, as assessed by individual staff development committees and an external data collector. A four-member team, consisting of a practicing master teacher, a practicing school administrator, a teacher educator from a similarly skilled subject area, and an independent evaluator from outside of the school (chosen from State Department of Education or the university) will serve as the evaluation team.

A number of instruments adapted from the Georgia Teacher Performance Assessment System are to be used to assess competence areas in the evaluation of new teachers.

To sum, the ATPE Basic Skills subtest was designed as a minimal competency based examination to measure and assess teacher proficiency in basic skill areas of reading, grammar, and mathematics.

The ATPE Professional knowledge was designed as a skill-based, criterion referenced examination. It was intended to measure specific and clearly defined classroom and knowledge skills required of the beginning teacher.

The teacher resident in Arizona will have received a Bachelor's Degree from an accredited institution and have completed an approved program in teacher education. He or she will possess an appropriate teaching certificate, passed the Basic Skills and Professional Knowledge subtests of the ATPE, and demonstrates proficiency in classroom skills during a resident program, before a permanent certification

decision is made.

Results of Basic Skills Testing of Arizona's Perspective Teachers

As mentioned earlier, the major focus of this study is to understand some of the background and impacts of basic skills testing of Arizona teachers. This section will look at the results of the ATPE Basic Skills component statewide, as well as the test results of students at Northern Arizona University.

Between July 1, 1981 and June 30, 1982, the Arizona Department of Education administered 1,903 ATPE Basic Skills exams. Of this total, 1,344 (60.6%) passed and 559 (29.4%) failed. Females performed slightly better than males with 1,016 out of 1,406 (73%) passed compared with 325 out of 488 (67%) passed. Minorities performed lower than Caucasian examinees to following extent:

ETHNIC GROUP	PASS	FAIL	TOTAL
Caucasian	1,252 (75%)	416 (25%)	1,668
Hispanics	43 (41%)	52 (59%)	105
Blacks	9 (26½%)	34 (73½%)	43
Native Americans	14 (26%)	40 (74%)	54

Of the 1,099 exams administered to graduates of out-of-state colleges or universities, 839 or approximately 76% were passed.

NAU Scores on the ATPE Basic Skills Component

According to data released by the Arizona Department of Education, for the period April 1 through June 30, 1982, 102 NAU students took the ATPE Basic Skills components of which 60 (58.8%) passed and 42 (41.2%) failed. Females performed slightly better better than males with 42 out of 75 (56%) passed compared with 14 out of 27 (51.9%) males passed. Minorities performed lower than Caucasian examinees to the following extent:

ETHNIC GROUP	PASS	FAIL	TOTAL
Caucasian	53 (66.3%)	27 (33.8%)	80
Hispanic	2 (28.6%)	5 (71.4%)	7
Native American	1 (10%)	9 (90%)	10
Asian	1 (50%)	1 (50%)	2

Looking at exam scores by certification field, elementary education majors had the highest pass rate with secondary the lowest. This is illustrated below:

CERTIFICATION FIELD	PASS	FAIL	TOTAL
Elementary	29 (64.4%)	16 (35.6%)	45
Secondary	23 (57.5%)	17 (42.5%)	40
Voc. Ed.	1 (100%)	0 (0%)	1
Special Ed.	7 (41.2%)	10 (58.8%)	17

Of the secondary fields, the major with the lowest

pass rate was physical education (7 out of 10 or 70% failed) and special education (10 out of 15 or 66.7% failed). Social studies (3), English (3), and Home Economics (2) had 100% passing rates.

Comparison of Results

One question that needs to be addressed is whether or not the data indicates a pattern of failure significantly different at NAU than at other teacher training institutions in the State. To date, the Arizona Department of Education has not released test results broken down by institution. It has given each school its own scores and also distributed summary material for the state as-a-whole, some of which has been previously discussed.

The purpose of a comparison of results is not to judge the quality of student or institution or level of instruction. On the contrary, the purpose is to suggest that without complete data sets, which account for incoming ability and other relevant variables, such comparisons are meaningless.

Take for example a recent NAU College of Education Newsletter which compares NAU scores on specific areas of the ATPE with the Statewide Average. The author concludes:

> It is obvious that the performance of the Northern Arizona University is significantly below the statewide average on a majority of the subskills and certainly on the total of the Arizona Teacher Proficiency Examination (The College Intercom, NAU College of Education, October 15, 1982).

Well there is really nothing obvious about the comparison at all. The number of students are not given, the significance

of the differences are not mentioned. All that is given is a subject area subskill, the NAU average, the statewide average and the difference as illustrated below:

ATPE READING

SUBSKILL	NAU	STATEWIDE	DIFFERENCE
Main Idea	94.3	94.3	0
Subordinate Detail	89.4	91.8	-2.4
Cause and Effect	88.2	87.9	+ .3
Sequencing Events	73.3	79.9	-6.6
Inferred Main idea	79.1	82.4	-3.3
Related Events	89.0	90.0	-1.0
Predict Outcomes	67.3	72.8	-5.5
Figurative Language	54.3	61.6	-7.3
Point of View	65.6	71.2	-5.6
Fact and Opinion	67.8	74.4	-6.6
Generalizations	84.0	87.2	-3.2
Contextual Meaning	70.0	72.1	-2.1
Synonyms	94.1	95.5	-1.4
Antonyms	80.0	85.7	-5.7
Homophones	98.8	98.3	+ .5
Alphabetic Order	95.0	95.1	- .1
Dictionary Skills	95.6	94.4	+1.2
Table of contents	86.2	91.4	-5.2
Read Maps, Charts, and Graphs	70.1	74.1	-4.0

What this kind of chart does is set up the conclusion (as is made by the author) that NAU students are "significantly below the statewide average..." and that the general public and prospective students will view this as a reflection on the caliber of education within the College.

What the chart looks like to us is a content analysis of the ATPE. But instead of using this data to establish validity of the Exam (or lack of it), the Department of Education releases this data so that the College of Education locates the blame for failure on itself. The legitimacy of the Examination goes unquestioned. The College of Education

begins to make changes so that they will not look bad by raising admissions standards and limiting access to the college for students who will do poorly on the Exam. Appearances take priority over program quality and the human experiences of those trying to become teachers, something that has typically been an upwardly mobile profession for students from lower class or working class backgrounds, are ignored.

Successes and Failures of NAU Students

As mentioned earlier, one of the responses by the College of Education at Northern Arizona University has been to institute an Admissions Exam (AE) for students wishing to enter the Teacher Education Program.

The Admissions Examination was based on the Arizona Teacher Proficiency Examination and members of the Math and English Departments along with College of Education faculty helped to make up the Exam. After revisions, the Exam consisted of three sections, reading, mathematics, and language arts, with 50 questions each.

To be admitted to the College, a student had to answer 70% of the Exam questions correctly. Additionally, a student must meet certain subject matter prerequisites and have a cumulative GPA of at least 2.5.

Data were collected from three separate testing dates, 10/81, 3/82, and 9/82, in order to ascertain what the Exam really showed and what other factors might be related to Exam success. In addition to Exam scores, students were asked to fill out a one page questionnaire concerning their

academic preparation. Finally, on one of the test dates, transcripts were obtained for students taking the Exam and academic achievement was compared with Exam success.

Test Results - October, 1981

The October, 1981, testing of perspective students presented a unique situation because it was the only time that students at a field-based site were tested. With one or two exceptions, field-based students were Navajo teacher aides enrolled in a federally subsidized program in which courses were taken on site so that teacher certification could be achieved.

In the October testing, 58 students were tested off campus while 83 students were tested on campus.

It was of interest to know what relationship might exist between college GPA and Admission Exam scores. The mean GPA for all students was 2.667 with a range from .545 to 4.00. Fifty-eight students (40.8%) had GPA's of less than 2.5, the minimum prerequisite for admission to the College of Education.

Table 1 shows a high positive correlation between GPA and Admission Exam scores.

TABLE 1

Pearson Product-Moment Correlations Between Student Grade Point Average and Scores on the Screening Examination.*

	MATH	VOCABULARY	READING COMP	SPELLING	USAGE	TOTAL
GPA	.5236	.6127	.5795	.5806	.5920	.6949
	N=141	N=140	N=141	N=141	N=141	N=141
	p<.001	p<.001	p<.001	p<.001	p<.001	p<.001

*This first version of the Admission Exam had five sections and was made up of 190 questions. This was later reduced to 150 questions in three sections as mentioned earlier.

In addition to being interested in the relationship between GPA and Exam score, one of the hopes of the College of Education was that increasing GPA requirements would reduce the number of failures on the Admission Exam (and on the Arizona Teacher Proficiency Exam).

To test this assumption, the sample was divided into two groups on the basis of GPA (2.5 or above and below 2.5). This was then cross-tabulated with Exam score (pass or fail). The results are represented below.

TABLE 2

GPA

ADMISSION EXAM	2.5 or Above	Below 2.5
PASS	71% (59)	34.4% (20)
FAIL	29% (24)	65.6% (38)
	100% (83)	100% (58)

N = 141

Chi Square = 18.565
df = 1
$p<.001$

Table 2 suggests that GPA is associated with success on the Admission Examination.

Also of interest was the difference in Exam success of students on campus and those at the field-based site. Whereas, 81% of those taking the Exam on campus passed, this was true for only 18.9% of those at the field-based

site. Since the vast majority of the latter were Navajo students, this suggests a failure rate of such magnitude to cause special concern for the Native American student.

One final interest concerned the extent of failure of students on campus with GPA's of at least 2.5 or above. Controlling for these variables, one finds that 15% of the students meeting these two criteria still failed the Admission Examination, a significant number of students.

Test Results - March, 1982

Between October, 1981, and April, 1982, the Admission Exam was given five times. Revisions were made in the Exam which reduced the number of questions to 150 in three sections: reading, mathematics, and language arts.

Thirty people were tested during the March period. In addition, each person was asked to fill out a one-page questionnaire concerning demographics and past academic performance.

Students were asked to evaluate their own ability in mathematics and English on a scale from 1 to 5 (A to F). The results are listed below.

SELF-EVALUATION IN MATH	SELF-EVALUATION IN ENGLISH
A = 7	A = 8
B = 14	B = 12
C = 8	C = 10
D = 1	

It was of interest to know whether these self-evaluations correlated significantly with student scores on the Admission

Examination. Table 3 indicates a high positive correlation between student self-evaluation and Admission Examination score.

TABLE 3

Pearson Product-Moment Correlations Between Student Self-Evaluations and Admission Exam Scores.

	MATH SCORE	READING SCORE	LANGUAGE ARTS SCORE
Self-Evaluation Math	.6617 N=30 p<.001	.5685 N=30 p<.002	.4620 N=30 p<.005
Self-Evaluation English	.4823 N=30 p<.003	.5753 N=30 p<.001	.4190 N=30 p<.02

Table 3 suggests two things which are important. First, students do seem to be able, with some degree of accuracy, to evaluate their own proficiency in English and mathematics. Given time and proper counseling or advisement, these students might well choose to seek the help or remediation necessary to pass the Admission Exam.

Second, student self-evaluation in math and math score had the highest correlation. Math proficiency, therefore, seems a good predictor of Exam success.

In terms of actual scores on the Admission Exam, 16.6% (5 out of 30) students failed to answer 70% of the questions correctly. Taken by section, four students failed the Reading section, six students failed the Math section, and four students failed the Language Arts section (failed to answer 70% of the questions in each section correctly).

Of the five students failing the Exam, four failed

all three parts of the Exam. One student failed the Math section with a score so low as to fail the entire Exam. Finally, one person passed the Exam but failed the Reading section if considered separately.

It would seem advisable for the College of Education to consider each section of the Exam separately, and make recommendations to students failing a particular section rather than consider the Exam as a whole. This is suggested even though the State considers only the total Exam score. For one thing this would give more selective advice to the student. For another thing, the State can at any time choose to consider each section of the ATPE separately.

One piece of information which might shed light on the separateness of each section of the Exam might be to look at the correlations between sections. This is presented in Table 4 below.

TABLE 4

Intercorrelation Among the Three Sections of the Admission Examination

	READING	MATHEMATICS	LANGUAGE ARTS
READING	---	.8610 P .001	.7544 P .001
MATHEMATICS	---	---	.8020 P .001
LANGUAGE ARTS	---	---	---

This suggests that the sections of the Admission Exam are significantly and positively correlated with each other. It also suggests high internal consistency and reliability

to the Exam.

Finally, it was of interest to know the relationship between ethnicity and Exam success. Broken down by ethnicity and Exam success, the results are as follows:

	PASS EXAM	FAIL EXAM
Caucasians	22	2
Native Americans	0	3
Hispanics	2	0
Orientals	1	0
	25	5

N = 30

Again, this indicates a pattern of minority failure to reach required levels of proficiency in basic skills as measured by the Admission Exam.

Test Results - September, 1982

In September, 1982, seventy-one freshmen and sophomore students enrolled in Introduction to Education were given the Admission Exam. Of this group, 25% were males and 75% were females, and the mean GPA was 2.921. Fifty-seven students identified themselves as Caucasian, ten as Mexican Americans or Spanish-speaking, and four as Native Americans.

Students were asked to report on their high school GPA and their self-report is listed below.

HIGH SCHOOL GPA

A = 18
B = 37
C = 15
D = 1
N = 71

High school average (self-reported) was then tabulated with the results of the Admission Exam with the following outcome:

ADMISSION EXAM			
PASS	FAIL		
18	0	A	HIGH SCHOOL GPA (SELF-REPORT)
34	3	B	
13	2	C	
0	1	D	

Chi Square = 12.962
df = 3
P .01

This would suggest that high school GPA presents a measure of student strength and weakness which is paralleled by the Admission Exam scores, and that students are able to accurately report on their own records. As expected, high GPA is associated with Exam success.

With regard to Exam success and ethnicity, the following pattern was found.

ADMISSION EXAM		
PASS	FAIL	
54	3	CAUCASIANS
3	1	NATIVE AMERICANS
8	2	HISPANICS
65	6	

This shows that six students (8.5%) failed the Admission Exam. This lower percentage (compared with previous testing

dates) may be attributed to the fact this was the first testing of the year and less able students might not have been as inclined to take the Exam at the first opportunity. The view is supported by the high mean GPA of the group (2.921). Still, of the six students that failed the exam, half were from minority ethnic backgrounds.

One other piece of information which might be used to predict student performance of the Admission Examination is ACT scores. From an earlier test date (September, 1981) ACT scores for thirty-five students were found and correlation coefficients computed using the ACT Academic Tests with the Admission Exam subsections. These results are presented in Table 5.

TABLE 5

Correlation Coefficient between scores on the Admission Examination and ACT Academic Tests

	ACT ACADEMIC TESTS			
	ENGLISH USAGE	MATHEMATICS USAGE	SOCIAL STUDIES READING	NATURAL SCIENCE READING
MATH	.4370 N=35 p<.005	.5890 N=35 p<.001	.2593 N=35 p<.05	.1550 N=35 p<.05
VOCABULARY	.6580 N=34 p<.001	.6053 N=34 p<.001	.4941 N=34 p<.001	.3005 N=34 p<.05
READING COMPREHENSION	.6289 N=34 p<.001	.5029 N=34 p<.001	.6004 N=34 p<.001	.3533 N=34 p<.02
SPELLING	.5062 N=34 p<.001	.2733 N=34 p<.05	.3000 N=34 p<.042	.2002 N=34 p<.05
USAGE	.5798 N=34 p<.001	.3314 N=34 p<.003	.2915 N=34 p<.047	.1881 N=34 p<.05
TOTAL	.7676 N=34 p<.001	.6096 N=34 p<.001	.5628 N=34 p<.001	.3537 N=34 p<.02

The results presented in Table 5 suggest that ACT scores and Admission Exam scores are highly correlated. In particular, the English and Mathematics Usage sections of the ACT are strongly associated with the Admission Exam Total scores.

Summary

The purpose of this research was to examine some of the impacts of basic skills proficiency testing in Arizona. What emerges from this work is an understanding of some of the promises, pitfalls, and pratfalls, associated with the Arizona experience.

Teacher testing, as a condition of certification, grew out of a perceived decline in the teacher applicant pool and a general concern with the quality of education. Teacher testing was one legislative response to these concerns.

The courts have ruled that testing must show "job-relatedness" and a "rational basis for use." In spite of high rates of minority failure, the courts have ruled that basic skills proficiency testing is related to the "legitimate employment needs" of the state. Court rulings on test validation have steered states towards criterion-referenced examinations which measure basic skills, professional knowledge, and content area ability.

Arizona has moved through a number of stages in the implementation of a teacher certification program which now includes a Teacher Proficiency Examination, Basic Skills and Professional Knowledge components. Also proposed is the use of a two year residency program now being piloted

in six districts in the State.

During the six month period beginning January 1, 1982, 1,903 Exams were administered by the Arizona Department of Education of which 1,344 (70.6%) passed the Basic Skills component at the required 80% proficiency level. Minorities performed at a much lower rate -- of the 193 Exams administered to blacks, Hispanics, and Native Americans, 66 (34.2%) passed.

Results for the field test period (50% correct responses) on the Professional Knowledge component show that 97% passed.

One response of the College of Education at Northern Arizona University has been to institute ar Admission Exam to screen applicants. One of the conclusions of this analysis is that large numbers of students are failing this Exam. Of the 242 student scores examined in this study, 73 (30%) failed at the required proficiency level, 70% of the questions answered correctly.

A second conclusion from the Admission Exam data is that large numbers of minority students are failing the Exam. Of the 78 blacks, Hispanics, and Native Americans tested, 53 (67.9%) failed. This has dismal implications for minority students and for the number of minority candidates completing teacher education at Northern Arizona University and qualifying for certification in Arizona.

In addition, high positive correlations were found among the scores on sections of the Exam as well as Exam scores and other measures of academic ability and achievement.

High school and college GPA, ACT scores, and student self-evaluation in mathematics and English were all significantly and positively associated with Exam scores.

The consistency of failure across so many different measures of ability and achievement suggest to this researcher that minority students may well find themselves penalized for their ethnic and cultural backgrounds. Without special help and remediation, a valuable resource, the representatives of minority cultures, will be lost to the schools of Arizona.

Finally, the high positive correlation between student self-evaluation and Exam scores indicates that students are able to recognize their own strengths and weakness in preparation and accomplishment.

Recommendations

Since students must take the Arizona Teacher Proficiency Examination anyway, and pass at the 80% proficiency level, it would make more sense to allow students to take the Basic Skills component as the Admission Exam rather than a separate, locally prepared, and possibly easier test. And although the State considers the ATPE as a whole (80% of the 150 questions must be answered correctly regardless of how one does on a particular section) it would provide greater diagnostic help if the grammar, reading comprehension, and mathematics sections were considered separately. Failure in one area could then lead to remediation.

The responsibility to provide necessary remediation must come as early as possible in the student's academic

career. The question of whose responsibility this is, the College of Education's or other academic divisions', should not be the issue.

It is recommended that the University form a center for the development of basic skills to meet the immediate needs of students and potential teachers. The ultimate relationship among high schools, community colleges, and universities in teaching basic skills, can then be addressed in reaching a long range solution.

Finally, advisement of undergraduate students is an area of key concern. Data must be collected on a regular basis from which informed decisions concerning proper sequence and level of coursework can be made. For example, if students score below required levels on the mathematics section of the Exam, they would be required to seek remediation before entering what are now the required math courses. This will require additional time and course requirements for the student. However, passing students from one course to the next, without the development of the required basic skills, serves neither the student nor the University.

REFERENCES

Anonymous. "Failing Teachers". THE MIAMI HERALD. January, 28, 1982.

_____. "Georgia Certification System". AEA ADVOCATE, 1982, 8(8), 19.

_____. "Results of Arizona Teacher Proficiency Test." AEA ADVOCATE, 1982, 8,8.

Arizona Department of Education. SYNOPSIS OF ORIGIN AND DEVELOPMENT - ARIZONA TEACHER PROFICIENCY EXAMINATION, 1981.

Brickell, H.M. "How to Change What Matters." EDUCATIONAL LEADERSHIP, 1980, 202-207.

Burke, F.G. "A Response to Shine and Goldman: Testing in New Jersey." EDUCATIONAL LEADERSHIP, 1980, 199-201.

Elam, S.M. "If Arizona's Shortage of Competent Teachers Grows Worse, Don't (Openly) Blame the Legislature." PHI DELTA KAPPAN, 1981, 63, 145-146.

Hathaway, W.E. "Testing Teachers." EDUCATIONAL LEADERSHIP, 1980, 38, 210-215.

Hazard, W.R., Freeman, L.D., Eisdorfer, S. and Tractenberg, P. LEGAL ISSUES IN TEACHER PREPARATION AND CERTIFICATION. Eric Clearinghous on Teacher Education: Washington, D.C., 1977.

Ivie, S.D. "Why Black Students Score Poorly on the NTE". HIGH SCHOOL JOURNAL, 1982, 65, 169-175.

Kelley, M.F. Report of: ARIZONA'S PERFORMANCE-BASED TEACHER CERTIFICATION PROGRAM, Arizona Department of Education, 1982.

Kelley, M.F. and Surbeck, E. Report of: PERFORMANCE-BASED TEACHER ASSESSMENT: A MODEL FOR LINKING EDUCATIONAL LEVELS, Arizona Department of Education, 1982.

McDaniel, T.R. "The NTE and Teacher Certification." PHI DELTA KAPPAN, 1977, 59, 186-188.

Medley, D.M. "The Role of Teacher Educators in Teacher Competency Assessment." INSERVICE: NATIONAL COUNCIL OF STATES ON INSERVICE EDUCATION, Syracuse University, New York, 1981, 9-12.

Merrit, R. "The Relationship of Performance in Professional Education Courses to national Teacher Examination Scores." EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT, 1980, 40, 1207-1209.

Nassif, P.M. and Solomon, L.M. THE STATE OF THE ART OF TEACHER CERTIFICATION TESTING. Symposium presented at the Annual Meeting of the National Council on Measurement in Education, National Evaluation Systems, Inc. New York, 1982.

Ornstein, A.C. "Accountability: Prospects for the 80's, SCHOOL AND COMMUNITY, 1981, 67, 24-25.

Pearson, C. "The Teacher Competency Movement: Blessing or Sham?" LEARNING, 1981, 10, 26-29.

Piper, M.K. and Houston, R.W. "The Search for Teacher Competence: CBTE and MCT." JOURNAL OF TEACHER EDUCATION. 1980. 31, (5), 37-40.

Piper, M.K. and O'Sullivan, P.S. "The National Teacher Examinations: Can It Predict Classroom Performance?" PHI DELTA KAPPAN, 1981, 401.

Popham, W.J. and Rankin, S.C. "Detroit's Measurement-Driven Instruction." EDUCATIONAL LEADERSHIP, 1980, 208-209.

Rubinstein, S.A., McDonough, M.W. and Allan, R.G. "The Changing Nature of Teacher Certification Programs" Paper presented at the Annual Meeting of the National Council on Measurement in Education, NATIONAL EVALUATION SYSTEMS, INC.: Amherst, MA. 1982.

Shine, W.A. and Goldman, N. "Goverance by Testing in New Jersey." EDUCATIONAL LEADERSHIP. 1980, 197-198.

Snowman, J., Leitner, D.W., Snyder, V. and Lockhard, L. "A Comparison of the Predictive Validities of Selected Academic Tests of the American College Test (ACT) Assessment Program and the Descriptive Tests of Language Skills for College Freshman in a Basic Skills Program." EDUCATIONAL AND PSYCHOLCGICAL MEASUREMENT, 1980, 40, 1159-1166.

Tractenberg, P.L. "Legal Issues in the Testing of School Personnels." PHI DELTA KAPPAN, 1979, 602-605.

Vorwerk, K.E. and Gorth, P.W. "Common Themes in Teacher Certification Testing Program Development and Implementation" NATIONAL EVALUATION SYSTEMS, INC.: Amherst, MA., 1982.

Weaver, W.T. "In Search of Quality: The Need for Talent in Teaching." PHI DELTA KAPPAN, 1979, 61, 29-32.

Webster, W.J. "Teacher Testing in Dallas." EDUCATIONAL LEADERSHIP, 1980, 216-217.

Willson, V.L. and Stoller J.E. "Predicting Teacher NTE Scores in Mathematics and Science." EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT, 1981, 41, 479-485.